

# PNEUMOCOCCAL PROTEIN HOMOLOGS AND FRAGMENTS FOR VACCINES

5

This application claims the priority of U.S. Provisional Application 60/150,750, filed August 25, 1999, the disclosure of which is hereby incorporated by reference in its entirety.

10

## FIELD OF THE INVENTION

15

This invention relates generally to the field of bacterial antigens and their use, for example, as immunogenic agents in humans and animals to stimulate an immune response. More specifically, it relates to the vaccination of mammalian species, especially humans, with one or more polypeptides derived from gram positive bacteria and which show sequence homology with an immunogenic polypeptide obtained from *Streptococcus pneumoniae*.

20

## BACKGROUND OF THE INVENTION

25

Polypeptides derived from gram positive bacteria are useful for stimulating production of antibodies that protect the vaccine recipient against infection by a wide range of serotypes of pathogenic gram positive bacteria, including *S. pneumoniae*. Further, the invention relates to antibodies against such polypeptides useful in diagnosis and passive immune therapy with respect to diagnosing and treating such pneumococcal infections.

30

The genus *Streptococcus* contains a variety of species responsible for causing disease in mammals, including humans, while also encompassing

005230-534960

species that constitute normal flora in humans and other mammals. Among the bacterial species implicated in the etiology of diseases in humans are *S. pyogenes* (part of the group A streptococcal bacteria, herein designated "GAS" for "group A streptococci"), *S. pneumoniae* (referred to as "pneumococcus") and *S. agalactiae* (the group B streptococci or "GBS"). The group A streptococci cause serious diseases such as necrotizing fasciitis, scarlet fever and sepsis, as well as less virulent diseases such as impetigo and pharyngitis. The pneumococci are the most common cause of community-acquired pneumonia and are also responsible for more than half of all cases of otitis media in children. The pneumococci are also the second most common pathogen associated with bacterial meningitis. The group B streptococci are the most prevalent pathogen associated with illness and death among newborns in the United States.

Currently, there are no vaccines available for the prevention of diseases caused by the group A and group B streptococci and presently available pneumococcal vaccines are not effective in children under 2 years of age or in the elderly due to the poor immunogenicity of the capsular carbohydrates that compose the current vaccine. It would therefore be highly advantageous to produce a vaccine that would prevent infection by these classes of pathogen, especially in the age groups mentioned.

In addition to the pathogens just described, some bacteria of the genus *Staphylococcus* are also of clinical importance. In fact, two of these are among the leading causes of nosocomial infections (infections acquired while in the hospital). Both *Staphylococcus aureus* and *Staphylococcus epidermidis* readily colonize the skin of healthy individuals and can cause acute disease in patients following immunosuppression or traumatic injury. Infections caused by these species include bacteremia, endocarditis, osteomyelitis, wound infections and infections associated with indwelling catheters.

*Streptococcus pneumoniae* is a gram positive bacterium that is a major causative agent in invasive infections in animals and humans, such as the aforementioned sepsis, meningitis, and otitis media, as well as lobar pneumonia (Tuomanen, *et al. New England J. of Medicine* **322**:1280-1284 (1995)). As part of the infection process, pneumococci readily bind to non-inflamed human epithelial cells of the upper and lower respiratory tract by binding to eukaryotic carbohydrates in a lectin-like manner (Cundell *et al., Micro. Path.* **17**:361-374 (1994)). Conversion to invasive pneumococcal infections for bound bacteria may involve the local generation of inflammatory factors which may activate the epithelial cells to change the number and type of receptors on their surface (Cundell, *et al., Nature*, **377**:435-438 (1995)). Apparently, one such receptor, platelet activating factor (PAF) is engaged by the pneumococcal bacteria and within a very short period of time (minutes) from the appearance of PAF, pneumococci exhibit strongly enhanced adherence and invasion of tissue. Certain soluble receptor analogs have been shown to prevent the progression of pneumococcal infections (Idanpaan-Heikkila *et al., J. Inf. Dis.*, **176**:704-712 (1997)). A number of other proteins have been suggested as being involved in the pathogenicity of *S. pneumoniae*.

20 *Streptococcus pneumoniae* itself has been shown to contain a gene  
which encodes a protein designated herein as Sp36. This protein has a  
predicted molecular mass of 91,538 Da and contains 5 histidine triad motifs  
(proposed to be involved in metal binding). The gene encoding this protein  
25 appears to be present in the 23 serotypes comprising the current commercially  
available pneumococcal-capsular vaccine. Immunization of mice with this  
protein, in the presence of Freund's adjuvant, stimulates an immune response  
which protects these mice from an intraperitoneal challenge with a dose of  
virulent pneumococci that would normally kill the mice.

For the reasons already stated above, there not only remains a need for identifying polypeptides having epitopes in common from various strains of *S. pneumoniae* but also from a broader spectrum of gram positive bacteria in order to utilize such polypeptides as vaccines to provide protection against a wide variety of infectious organisms.

### BRIEF SUMMARY OF THE INVENTION

In accordance with the present invention, there is provided vaccines that include polypeptides obtained from gram positive bacteria other than *S. pneumoniae*, as well as variants of said polypeptides and active fragments of such polypeptides.

The present invention is also directed to novel genes, and the polypeptides encoded thereby, derived from gram positive bacteria other than *S. pneumoniae*, and which bear sequence homology to the Sp36 gene already described. Such gram positive bacteria include the group A and B streptococci, as described herein, as well as species of the genus *Staphylococcus*, especially *S. aureus*.

In a particular embodiment, the present invention is directed to specific gene sequences, and proteins encoded thereby, derived from the group A and group B streptococci, and to the use of such expressed polypeptides and proteins as the basis for pharmaceutical compositions useful as vaccines and as a means for enabling isolation of antibodies with therapeutic and/or prophylactic activity (such as would be useful in preparing products like CytoGam).

In a further embodiment, the present invention also relates to the preparation and use of fragments of the novel polypeptides disclosed herein,

such fragments being immunogenic in nature and being useful in the preparation of vaccines against diseases caused by the pathogens from which such polypeptides, and fragments thereof, are derived.

5

### BRIEF DESCRIPTION OF DRAWINGS

10

Figures 1 shows the results of a Southern blot of genomic DNA from *S. aureus*, *S. pyogenes*, and pneumococcus. The DNA was digested with restriction nucleases *Bam*HI or *Pvu*II, and after electrophoresis and transfer to a nylon membrane, was probed with a labeled DNA fragment encompassing the pneumococcal gene encoding Sp36. The hybridization and washes were carried out under low stringency conditions. The results show hybridization by the labeled probe to a *S. aureus* fragment in both the *Bam*HI and *Pvu*II lanes and to two fragments in the *Pvu*II digests of two strains of *S. pyogenes*.

20

*C. Suh* Figures 2 shows an alignment between the Sp36 amino acid sequence from *S. pneumoniae* strain N4 and the homologous sequences from *S. pyogenes* and *S. agalactiae*. Amino acids identical to those of the polypeptide from *S. pneumoniae* are boxed.

25

Figure 3 shows the results of a Southern blot of genomic DNA from *S. pyogenes*, *S. agalactiae*, and *S. pneumoniae* probed with DNA encoding the full length Sp36 homolog from *S. pyogenes*. The hybridization was carried out under low stringency conditions. These results demonstrate that the *S.*

pyogenes Sp36 homolog, used as a probe, is capable of detecting a homologous gene in *S. agalactiae* and pneumococcus.

Figure 4 shows the results of a western blot using rabbit polyclonal antiserum generated against recombinant Sp36 protein cloned from *S. pneumoniae* strain Norway 4. The results demonstrate that this antiserum not only reacts with the protein against which it was raised (here, Sp36), as well as to a protein of similar size in a lysate of a serotype 6B strain of pneumococcus, but also reacts with a recombinant protein encoded by the Sp36 homolog gene of group B streptococci.

Sub Co Figure 5 shows the amino acid sequence for the GAS36 homologs with the histidine triad regions underlined (Fig. 5(a) and (b)) and the sequence for a GBS36 homolog (Fig. 5(c)) with its histidine triad regions underlined.

#### DETAILED DESCRIPTION OF THE INVENTION

The present invention is directed to novel polynucleotides and polypeptides derived from species of gram positive bacteria, especially group A and B streptococci, and including the genus *Staphylococcus*, most especially *S. pyogenes* (GAS), *S. agalactiae* (GBS), and *S. aureus*, respectively.

Further, the present invention is directed to polynucleotides derived from gram positive bacteria and which are at least partially homologous to the polynucleotides making up the gene coding for the previously disclosed Sp36 gene of *S. pneumoniae* (U.S. Application Serial No. 60/113,048).



The present invention is also directed to polynucleotides, and immunologically active fragments, segments, or portions, thereof, which polypeptides are encoded by the polynucleotides disclosed herein.

5 The present invention also relates to such polynucleotides and polypeptides in enriched, preferably isolated, or even purified, form.

10 In accordance with the present invention, the term "DNA segment" refers to a DNA polymer, in the form of a separate fragment or as a component of a larger DNA construct, which has been derived from DNA isolated at least once in substantially pure form, i.e., free of contaminating endogenous materials and in a quantity or concentration enabling identification, manipulation, and recovery of the segment and its component nucleotide sequences by standard biochemical methods, for example, using a cloning vector. Such segments are provided in the form of an open reading frame uninterrupted by internal nontranslated sequences, or introns, which are typically present in eukaryotic genes. Sequences of non-translated DNA may be present downstream from the open reading frame, where they do not interfere with manipulation or expression of the coding regions.

20 The nucleic acids and polypeptide expression products disclosed according to the present invention, as well as expression vectors containing such nucleic acids and/or such polypeptides, may be in "enriched form." As used herein, the term "enriched" means that the concentration of the material is at least about 2, 5, 10, 100, or 1000 times its natural concentration (for example), advantageously 0.01%, by weight, preferably at least about 0.1% by weight. Enriched preparations of about 0.5%, 1%, 5%, 10%, and 20% by weight are also contemplated. The sequences, constructs, vectors, clones, and other materials comprising the present invention can advantageously be in enriched or isolated form.

25 30

"Isolated" in the context of the present invention with respect to polypeptides (or polynucleotides) means that the material is removed from its original environment (*e.g.*, the natural environment if it is naturally occurring).

5 For example, a naturally-occurring polynucleotide or polypeptide present in a living organism is not isolated, but the same polynucleotide or polypeptide, separated from some or all of the co-existing materials in the natural system, is isolated. Such polynucleotides could be part of a vector and/or such polynucleotides or polypeptides could be part of a composition, and still be  
10 isolated in that such vector or composition is not part of its natural environment. The polypeptides and polynucleotides of the present invention are preferably provided in an isolated form, and most preferably are purified to homogeneity.

15 The polynucleotides, and recombinant or immunogenic polypeptides, disclosed in accordance with the present invention may also be in "purified" form. The term "purified" does not require absolute purity; rather, it is intended as a relative definition, and can include preparations that are highly purified or preparations that are only partially purified, as those terms are  
20 understood by those of skill in the relevant art. For example, individual clones isolated from a cDNA library have been conventionally purified to electrophoretic homogeneity. Purification of starting material or natural material to at least one order of magnitude, preferably two or three orders, and more preferably four or five orders of magnitude is expressly  
25 contemplated. Furthermore, claimed polypeptides having a purity of preferably 0.001%, or at least 0.01% or 0.1%; and even 1% by weight or greater is expressly contemplated.

The term "coding region" refers to that portion of a gene which either  
30 naturally or normally codes for the expression product of that gene in its



006380" 53354960

natural genomic environment, i.e., the region coding *in vivo* for the native expression product of the gene. The coding region can be from a normal, mutated or altered gene, or can even be from a DNA sequence, or gene, wholly synthesized in the laboratory using methods well known to those of skill in the art of DNA synthesis.

In accordance with the present invention, the term "nucleotide sequence" refers to a heteropolymer of deoxyribonucleotides. Generally, DNA segments encoding the proteins provided by this invention are assembled from cDNA fragments and short oligonucleotide linkers, or from a series of oligonucleotides, to provide a synthetic gene which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon.

The term "expression product" means that polypeptide or protein that is the natural translation product of the gene and any nucleic acid sequence coding equivalents resulting from genetic code degeneracy and thus coding for the same amino acid(s).

The term "fragment," when referring to a coding sequence, means a portion of DNA comprising less than the complete coding region whose expression product retains essentially the same biological function or activity as the expression product of the complete coding region.

The term "primer" means a short nucleic acid sequence that is paired with one strand of DNA and provides a free 3'OH end at which a DNA polymerase starts synthesis of a deoxyribonucleotide chain.

The term "promoter" means a region of DNA involved in binding of RNA polymerase to initiate transcription.

The term "open reading frame (ORF)" means a series of triplets coding for amino acids without any termination codons and is a sequence (potentially) translatable into protein.

5

As used herein, reference to a DNA sequence includes both single stranded and double stranded DNA. Thus, the specific sequence, unless the context indicates otherwise, refers to the single strand DNA of such sequence, the duplex of such sequence with its complement (double  
10 stranded DNA) and the complement of such sequence.

In accordance with the present invention, the term "percent identity" or "percent identical," when referring to a sequence, means that a sequence is compared to a claimed or described sequence after alignment of the sequence  
15 to be compared (the "Compared Sequence") with the described or claimed sequence (the "Reference Sequence"). The Percent Identity is then determined according to the following formula:

$$\text{Percent Identity} = 100 [1 - (C/R)]$$

20

wherein C is the number of differences between the Reference Sequence and the Compared Sequence over the length of alignment between the Reference Sequence and the Compared Sequence wherein (i) each base or amino acid in  
25 the Reference Sequence that does not have a corresponding aligned base or amino acid in the Compared Sequence and (ii) each gap in the Reference Sequence and (iii) each aligned base or amino acid in the Reference Sequence that is different from an aligned base or amino acid in the Compared Sequence, constitutes a difference; and R is the number of bases or amino acids in the  
30 Reference Sequence over the length of the alignment with the Compared

Sequence with any gap created in the Reference Sequence also being counted as a base or amino acid.

5 If an alignment exists between the Compared Sequence and the Reference Sequence for which the percent identity as calculated above is about equal to or greater than a specified minimum Percent Identity then the Compared Sequence has the specified minimum percent identity to the Reference Sequence even though alignments may exist in which the hereinabove calculated Percent Identity is less than the specified Percent  
10 Identity.

Thus, the present invention is directed to novel, isolated polypeptides comprising an amino acid sequence at least 75% identical to a sequence in SEQ ID NO: 2, 4 or 6, preferably polypeptides at least 90% identical thereto,  
15 more preferably 95% identical to the sequence of SEQ ID NO: 2 or 4, and most preferably having the sequence of either SEQ ID NO: 2 or 4.

The isolated polypeptides of the present invention may be found in a wide variety of microorganisms, but will commonly be found in an organism  
20 selected from the group consisting of group A streptococci, group B streptococci, and *Staphylococcus aureus*, and wherein the group A streptococcal organism is *Streptococcus pyogenes* and the group B streptococcal organism is *Streptococcus agalactiae*. Also, polypeptides of the invention include, but are in no way limited to, isolated polypeptides having a  
25 sequence at least 25% identical to the amino acid sequence of the Sp36 protein of *Streptococcus pneumoniae*.

The present invention further relates to immunogenically active fragments of the isolated polypeptides disclosed herein.

005230 "SEQ ID NO:2, 4 or 6" 0054360

The terms "fragment," "derivative" and "analog" when referring to the polypeptides disclosed herein means a polypeptide which retains essentially the same biological function or activity as such polypeptide. Thus, an analog includes a proprotein, or preprotein, which can be activated by cleavage of the proprotein portion to produce an active mature polypeptide. Such fragments, derivatives and analogs must have sufficient similarity to the polypeptide of SEQ ID NO:2, 4 or 6 so that immunogenic activity of the native polypeptide is retained.

The polypeptide of the present invention may be a recombinant polypeptide, a natural polypeptide or a synthetic polypeptide, preferably a recombinant polypeptide.

The fragment, derivative or analog of the polypeptide of SEQ ID NO:2, 4, or 6 may be (i) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the mature polypeptide is fused with another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or (iv) one in which the additional amino acids are fused to the mature polypeptide, such as a leader or secretory sequence or a sequence which is employed for purification of the mature polypeptide or a proprotein sequence. Such fragments, derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.

As known in the art "similarity" between two polypeptides is determined by comparing the amino acid sequence and its conserved amino acid substitutes of one polypeptide to the sequence of a second polypeptide.

Fragments or portions of the polypeptides of the present invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, the fragments may be employed as intermediates for producing the full-length polypeptides. Fragments or portions of the polynucleotides of the present invention may be used to synthesize full-length polynucleotides of the present invention.

As used herein with reference to polypeptides, the terms "portion," "segment," and "fragment," refer also to a continuous sequence of residues, such as amino acid residues, which sequence forms a subset of a larger sequence. For example, if a polypeptide were subjected to treatment with any of the common endopeptidases, such as trypsin, chymotrypsin, or papain, the oligopeptides resulting from such treatment would represent portions, segments or fragments of the starting polypeptide.

The present invention is also directed to isolated polynucleotides whose sequences contain coding regions encoding the polypeptides of the present invention, preferably the polypeptides of SEQ ID NO: 2, 4, and 6 and most preferably will be the isolated polynucleotides comprising the sequences of SEQ ID NOS: 1, 3, and 5.

The present invention is also directed to fragments or portions of such sequences which contain at least 15 bases, preferably at least 30 bases, more preferably at least 50 bases and most preferably at least 80 bases, and to those sequences which are at least 60%, preferably at least 80%, and most preferably at least 95%, especially 98%, identical thereto, and to DNA (or RNA) sequences encoding the same polypeptide as the sequences of SEQ ID NOS: 2, 4, and 6, including fragments and portions thereof and, when derived from natural sources, includes alleles thereof.

30



Yet another aspect of the present invention is directed to an isolated DNA (or RNA) sequence or molecule comprising at least the coding region of a bacterial gene (or a DNA sequence encoding the same polypeptide as such coding region), in particular an expressed bacterial gene, which bacterial gene comprises a DNA sequence homologous with, or contributing to, the sequence depicted in SEQ ID NOS: 1, 3, and 5 or one at least 60%, preferably at least 80%, and most preferably at least 95%, especially 98%, identical thereto, including 100% identity, as well as fragments or portions of the coding region which encode a polypeptide having a similar function to the polypeptide encoded by said coding region. Thus, the isolated DNA (or RNA) sequence may include only the coding region of the expressed gene (or fragment or portion thereof as hereinabove indicated) or may further include all or a portion of the non-coding DNA (or RNA) of the expressed bacterial gene.

In general, sequences homologous with and contributing to the sequences of SEQ ID NOS: 1, 3, and 5 (or one at least 60%, preferably at least 80%, and most preferably at least 95% identical or homologous thereto) are from the coding region of a bacterial gene.

The polynucleotides according to the present invention may also occur in the form of mixtures of polynucleotides hybridizable to some extent with the gene sequences containing any of the nucleotide sequences of SEQ ID NOS: 1, 3, and 5, including any and all fragments thereof, and which polynucleotide mixtures may be composed of any number of such polynucleotides, or fragments thereof, including mixtures having at least 10, perhaps at least 30 such sequences, or fragments thereof.

Fragments of the full length polynucleotide of the present invention may be used as hybridization probes for a DNA library to isolate the full length DNA

and to isolate other DNAs which have a high sequence similarity to the gene or similar biological activity. Probes of this type preferably have at least 15 bases, may have at least 30 bases and even 50 or more bases. The probe may also be used to identify a DNA clone corresponding to a full length transcript and a genomic clone or clones that contain the complete gene including regulatory and promotor regions. An example of a screen comprises isolating the coding region of the gene by using the known DNA sequence to synthesize an oligonucleotide probe. Labeled oligonucleotides having a sequence complementary to that of the gene of the present invention are used to screen a library of DNA or mRNA to determine which members of the library the probe hybridizes to.

The present invention is also directed to vectors comprising the polynucleotides disclosed herein, as well as to genetically engineered cells comprising such vectors and/or polynucleotides. Thus, the present invention also relates to vectors which include polynucleotides of the present invention, host cells which are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques.

Host cells are genetically engineered (transduced or transformed or transfected) with the vectors of this invention which may be, for example, a cloning vector or an expression vector. The vector may be, for example, in the form of a plasmid, a viral particle, a phage, etc. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the genes of the present invention. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

005230 "GEB-960

The polynucleotides of the present invention may be employed for producing polypeptides by recombinant techniques. Thus, for example, the polynucleotide may be included in any one of a variety of expression vectors for expressing a polypeptide. Such vectors include chromosomal, nonchromosomal and synthetic DNA sequences, e.g., derivatives of SV40; bacterial plasmids; phage DNA; baculovirus; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies. However, any other vector may be used as long as it is replicable and viable in the host.

10

The appropriate DNA sequence may be inserted into the vector by a variety of procedures. In general, the DNA sequence is inserted into an appropriate restriction endonuclease site(s) by procedures known in the art. Such procedures and others are deemed to be within the scope of those skilled in the art.

15

The DNA sequence in the expression vector is operatively linked to an appropriate expression control sequence(s) (promoter) to direct mRNA synthesis. As representative examples of such promoters, there may be mentioned: LTR or SV40 promoter, the E. coli. lac or trp, the phage lambda P<sub>L</sub> promoter and other promoters known to control expression of genes in prokaryotic or eukaryotic cells or their viruses. The expression vector also contains a ribosome binding site for translation initiation and a transcription terminator. The vector may also include appropriate sequences for amplifying expression.

20  
25

In addition, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or neomycin resistance

for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in E. coli.

5 The vector containing the appropriate DNA sequence as hereinabove described, as well as an appropriate promoter or control sequence, may be employed to transform an appropriate host to permit the host to express the protein.

10 As representative examples of appropriate hosts, there may be mentioned: bacterial cells, such as E. coli, Streptomyces, Salmonella typhimurium; fungal cells, such as yeast; insect cells such as Drosophila S2 and Spodoptera Sf9; animal cells such as CHO, COS or Bowes melanoma; adenoviruses; plant cells, etc. The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein.

15 More particularly, the present invention also includes recombinant constructs comprising one or more of the sequences as broadly described above. The constructs comprise a vector, such as a plasmid or viral vector, into which a sequence of the invention has been inserted, in a forward or  
20 reverse orientation. In a preferred aspect of this embodiment, the construct further comprises regulatory sequences, including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available. The following vectors are provided by way of example; Bacterial:  
25 pQE70, pQE60, pQE-9 (Qiagen), pBS, pD10, phagescript, phiX174, pBluescript SK, pBSKS, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene); pTRC99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia); Eukaryotic: pWLNEO, pSV2CAT, pOG44, pXT1, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia). However, any other plasmid or vector may be used as long as  
30 they are replicable and viable in the host.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda P<sub>R</sub>, P<sub>L</sub> and trp. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art.

10

In a further embodiment, the present invention relates to host cells containing the above-described constructs. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-Dextran mediated transfection, or electroporation (Davis, L., Dibner, M., Battey, I., Basic Methods in Molecular Biology, (1986)).

20

The constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. Alternatively, the polypeptides of the invention can be synthetically produced by conventional peptide synthesizers.

25

Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., Molecular Cloning: A Laboratory Manual, Second Edition,

30



Cold Spring Harbor, N.Y., (1989), the disclosure of which is hereby incorporated by reference.

Transcription of the DNA encoding the polypeptides of the present invention by higher eukaryotes is increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp that act on a promoter to increase its transcription. Examples include the SV40 enhancer on the late side of the replication origin bp 100 to 270, a cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of E. coli and S. cerevisiae Trp1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK),  $\alpha$ -factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an N-terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product.

Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic

selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include E. coli, Bacillus subtilis, Salmonella typhimurium and various species within the genera Pseudomonas,  
5 Streptomyces, and Staphylococcus, although others may also be employed as a matter of choice.

As a representative but nonlimiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of  
10 replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM1 (Promega Biotec, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate  
15 promoter and the structural sequence to be expressed.

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced by appropriate means (e.g., temperature shift or chemical induction) and cells are  
20 cultured for an additional period.

Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further  
25 purification.

Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents, such methods are well known to those  
30 skilled in the art.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell, 23:175 (1981), and other cell lines capable of expressing a compatible  
5 vector, for example, the C127, 3T3, CHO, HeLa and BHK cell lines. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional  
10 termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements.

The polypeptide can be recovered and purified from recombinant cell cultures by methods including ammonium sulfate or ethanol precipitation, acid  
15 extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC)  
20 can be employed for final purification steps.

The polypeptides of the present invention may be a naturally purified product, or a product of chemical synthetic procedures, or produced by recombinant techniques from a prokaryotic or eukaryotic host (for example, by  
25 bacterial, yeast, higher plant, insect and mammalian cells in culture). Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. Polypeptides of the invention may also include an initial methionine amino acid residue.

5 The polypeptides of the present invention, when utilized for clinically related purposes, may also be suspended in a pharmacologically acceptable diluent or excipient to facilitate such uses, which will include use as a vaccine for the purpose of preventing a wide variety of streptococcal and staphylococcal infections.

10 In accordance with another aspect of the present invention, there is provided a vaccine that includes at least one polypeptide that is at least 75% identical, preferably at least 90% identical and most preferably 95% identical, to a polypeptide sequence comprising the sequence of SEQ ID NO: 2, 4, or 6. Such variations in homology for putative vaccines are well known in the art (See, for example, Hanson et al., "Active and Passive Immunity Against *Borrelia burgdorferi* Decorin Binding Protein A (DbpA)," Infection and Immunity, (May) 1998, p. 2143 - 2153; Roberts et al., "Heterogeneity Among Genes Including Decorin Binding Proteins A and B of *Borrelia burgdorferi sensu lato*," Infection and Immunity, (Nov) 1998, p. 5275-5285). Such observations would similarly apply to portions, segments or fragments of the polypeptides disclosed herein.

20 Such segments find a multitude of uses. For example, such segments of the polypeptides according to the present invention find use as intermediates in the synthesis of higher molecular weight structures also within the present invention.

25 The term "active fragment" means a fragment that generates an immune response (i.e., has immunogenic activity) when administered, alone or optionally with a suitable adjuvant, to an animal, such as a mammal, for example, a rabbit or a mouse, and also including a human.

In accordance with a further aspect of the invention, a vaccine of the type hereinabove described is administered for the purpose of preventing or treating infection caused by streptococci and staphylococci as well as many related organisms.

5

A vaccine in accordance with the present invention may include one or more of the hereinabove described polypeptides or active fragments thereof. When employing more than one polypeptide or active fragment, such as two or more polypeptides and/or active fragments may be used as a physical mixture  
10 or as a fusion of two or more polypeptides or active fragments. The fusion fragment or fusion polypeptide may be produced, for example, by recombinant techniques or by the use of appropriate linkers for fusing previously prepared polypeptides or active fragments.

15

In many cases, the variation in the polypeptide or active fragment is a conservative amino acid substitution, although other substitutions are within the scope of the invention.

20

In accordance with the present invention, a polypeptide variant includes variants in which one or more amino acids are substituted and/or deleted and/or inserted.

25

In another aspect, the invention relates to passive immunity vaccines formulated from antibodies against a polypeptide or active fragment of a polypeptide of the present invention. Such passive immunity vaccines can be utilized to prevent and/or treat streptococcal and staphylococcal infections in patients. In this manner, according to a further aspect of the invention, a vaccine can be produced from a synthetic or recombinant polypeptide of the present invention or an antibody against such polypeptide.

30



Still another aspect the present invention relates to a method of using one or more antibodies (monoclonal, polyclonal or sera) to the polypeptides of the invention as described above for the prophylaxis and/or treatment of diseases that are caused by streptococcal and staphylococcal bacteria. In particular, the invention relates to a method for the prophylaxis and/or treatment of infectious diseases that are caused by streptococci and staphylococci. In a still further preferred aspect, the invention relates to a method for the prophylaxis and/or treatment of such diseases as necrotizing fasciitis, scarlet fever, sepsis and many diseases of newborns, in humans by utilizing a vaccine of the present invention.

Generally, vaccines are prepared as injectables, in the form of aqueous solutions or suspensions. Vaccines in an oil base are also well known such as for inhaling. Solid forms which are dissolved or suspended prior to use may also be formulated. Pharmaceutical carriers, diluents and excipients are generally added that are compatible with the active ingredients and acceptable for pharmaceutical use. Examples of such carriers include, but are not limited to, water, saline solutions, dextrose, or glycerol. Combinations of carriers may also be used.

Vaccine compositions may further incorporate additional substances to stabilize pH, or to function as adjuvants, wetting agents, or emulsifying agents, which can serve to improve the effectiveness of the vaccine.

Vaccines are generally formulated for parenteral administration and are injected either subcutaneously or intramuscularly. Such vaccines can also be formulated as suppositories or for oral administration, using methods known in the art, or for administration through nasal or respiratory routes.

005280" 6E851950

The amount of vaccine sufficient to confer immunity to pathogenic bacteria is determined by methods well known to those skilled in the art. This quantity will be determined based upon the characteristics of the vaccine recipient and the level of immunity required. Typically, the amount of vaccine to be administered will be determined based upon the judgment of a skilled physician. Where vaccines are administered by subcutaneous or intramuscular injection, a range of 0.5 to 500  $\mu$ g purified protein may be given.

The present invention is also directed to a vaccine in which a polypeptide or active fragment of the present invention is delivered or administered in the form of a polynucleotide encoding the polypeptide or active fragment, whereby the polypeptide or active fragment is produced *in vivo*. The polynucleotide may be included in a suitable expression vector and combined with a pharmaceutically acceptable carrier.

Thus, the present invention expressly contemplates a vaccine composition comprising any of the polypeptides disclosed herein, said polypeptide being present in an amount effective to produce an immune response, and wherein said polypeptide is suspended in a pharmacologically acceptable carrier, diluent or excipient.

The vaccine compositions of the present invention may also comprise live vaccines, containing such organisms as *Streptococcus gordoniae* and *Salmonella typhi*, wherein said organisms contain recombinant polypeptides as disclosed herein.

In addition, the polypeptides of the present invention can be used as immunogens to stimulate the production of antibodies for use in passive immunotherapy, for use as diagnostic reagents, and for use as reagents in other processes such as affinity chromatography.

Thus, the present invention is also directed to methods for the prevention of a wide variety of diseases caused by streptococcal and staphylococcal organisms, said methods involving the administering of  
5 vaccines disclosed herein to animals at risk of such diseases, especially where said animals are humans.

In addition, the invention disclosed herein is also directed to a means of treating animals, especially humans, afflicted with a disease caused by the  
10 organisms from which the isolated polypeptides of the invention are derived, such methods including, but not being limited to, administering to an animal, especially a human, afflicted with such a disease of a therapeutically effective amount of an antibody, or mixture of antibodies, against the polypeptides disclosed herein.

15 Antibodies specific for the polypeptides disclosed herein may be either polyclonal or monoclonal and may even be in the form of antisera. When such antibodies are monoclonal in nature, they may be produced by conventional methods of preparing monoclonal antibodies, such as from  
20 conventional hybridoma cells, and may also be produced by genetically engineered cells transformed with vectors containing genes specifically coding for the different heavy and light chains of antibody molecules having an arrangement of variable regions specifically complementary to one or more of the polypeptides of the invention. Such recombinantly produced  
25 antibodies may be in the form of either dimers or tetramers, depending on the type of cellular expression system utilized therefor.

The invention will now be further described in more detail in the following non-limiting examples and it will be appreciated that additional and  
30 different embodiments of the teachings of the present invention will

doubtless suggest themselves to those of skill in the art and such other embodiments are considered to have been inferred from the disclosure herein.

5

### Example 1

#### 10 Southern Blot Analysis of Chromosomal DNA Using Probes Specific for the Sp36 Gene of *Streptococcus pneumoniae*

Genomic DNA was isolated from *Staphylococcus aureus*, *Streptococcus pyogenes* (group A), and *Streptococcus agalactiae* (group B) after overnight growth of the bacteria. The DNA was digested to completion by overnight incubation with restriction enzymes (*Bam*HI and *Pvu*II), and then DNA fragments were resolved by size by agarose gel electrophoresis before transfer to a nylon membrane. The membrane was then probed with DNA encoding the entire Sp36 open reading frame that had been fluorescein-labeled with random primers using a kit from Amersham Pharmacia Biotech Inc. The hybridization and washes were carried out under low stringency conditions (i.e., 45°C, 5xSSC hybridization; 45°C, 1xSSC for 1<sup>st</sup> wash; 45°C, 0.5xSSC for 2<sup>nd</sup> wash). Here, SSC is composed of 150 mM NaCl and 15 mM sodium citrate, pH 7.0 and all washes are 50 mL each.

25

After hybridization and washing was complete, the bound, fluorescein-labeled probe was detected using an anti-fluorescein antibody as per the manufacturer's instructions with the kit. Similarly digested DNA from *Streptococcus pneumoniae* strain SJ2 (serotype 6B) was used as a positive

control. Fluorescein-labeled bacteriophage lambda DNA digested with the restriction nuclease *Hind*III was used as a size marker.

The Sp36 probe hybridized with a single fragment in the digested *S. aureus* DNA (~4.5 kb *Bam*HI fragment, ~5 kb *Pvu*II fragment) and with 2 major fragments in a *Pvu*II digest of serotype M1 of the group A streptococci genomic DNA (~4.0 kb, and ~4.2 kb ).

## Example 2

10

### BLAST Analysis Using Sp36 Predicted Amino Acid Sequence

15 ~~Subtil~~ Sequence comparisons of the Sp36 encoded protein sequence against the publicly available GenBank sequence database (including the unfinished database microbial (<http://www.ncbi.nlm.nih.gov/BLAST/unfinishedgenome.html>)) revealed two highly homologous amino acid sequences. One of these was a predicted amino acid sequence from the *S. pyogenes* genome. This predicted polypeptide comprised 825 amino acid residues (MW = 92,616 Da) that was 25.1% identical to the Sp36 amino acid sequence from pneumococcus serotype 4 but maintained the 5 histidine triads (underlined in Figure 5(a) - SEQ ID NO: 2). The second polypeptide encoded within the *S. pyogenes* database contained several errors that were corrected by our sequencing of this region of the genome. The DNA fragment obtained encoded a protein of 792 amino acids (MW = 87,457 Da) that was 12.6% identical to the pneumococcal sequence and 12.5% identical to the first *S. pyogenes* polypeptide. This predicted amino acid sequence contained four histidine triad motifs (underlined in Fig. 5(b) - SEQ ID NO.: 4). The third polypeptide sequence obtained was one already in the database (Accession No. AF062533) and identified only as an unknown gene downstream from a gene identified as *Imb* in *S. galactiae*. This 822

20  
25  
30



amino acid protein thus has a predicted molecular weight of 92,353 Da and maintains the 5 histidine triad motifs (underlined in Figure 5(C) - SEQ ID NO: 6). This second polypeptide shows 25.6% sequence identity to Sp36 of pneumococcus type 4 and 97.7% and 11.6% identity to the two group A homologs, respectively.

### Example 3

10

#### **Southern Blot Analysis Using a group A Streptococcal Sp36 Homolog Probe**

005230" 46354960  
Southern blot analysis was performed with a fluorescein-labeled DNA fragment as probe, which encoding a group A streptococcal Sp36 homolog cloned from an M1 serotype of the group A streptococcal genome. This fragment was then used to probe genomic DNA from an M6 serotype of the group A streptococcal genome, as well as serotype 1a and serotype 3 of the group B streptococcal genome, and strain SJ2 (serotype 6B) of pneumococcus. In all cases, a single band was obtained in DNA digested with *Bam*HI when hybridization was carried out under low stringency conditions (as described above). A band of about 20 kb was visualized in group A streptococcal DNA, about 4.5 kb was obtained for group B streptococcal DNA, and a band of about 4kb was seen for pneumococcus.

25

### Example 4

#### **Western Blot Analysis of Reactivity of group B Streptococcal Homolog With Anti-Pneumococcal Sp36 Antiserum**

30

005280 "54433" 08500

To determine whether antiserum raised against recombinant Sp36 from *S. pneumoniae* would recognize the recombinant Sp36 homolog encoded by group B streptococcal organisms, a western blot was performed. One hundred nanograms (100 ng) of recombinant Sp36 polypeptide cloned from either *S. pneumoniae* serotype 4, or of the Sp36 homolog cloned from group B streptococcal organisms, or from an unrelated recombinant protein control expressed and purified in the same way, were subjected to SDS-PAGE containing 12% acrylamide. A cell lysate of pneumococcal strain SJ2 (serotype 6B) was also included on the gel. After electrophoresis, the separated proteins were transferred to a nitrocellulose membrane and probed with rabbit polyclonal antiserum raised against the recombinant pneumococcal protein. Bound antibodies were detected chemiluminescently with a goat anti-rabbit IgG antibody conjugated to horseradish peroxidase using the substrate ECL (from Amersham). The results demonstrate that antiserum raised against the pneumococcal Sp36 protein cross-react with the Sp36 homolog identified from the group B streptococci and thereby indicating conservation of epitopes between the proteins. The group B streptococcal homolog is also approximately the same size as the protein detected in *S. pneumoniae* lysates. Because the group A and B homologs are highly homologous, if not identical, such antiserum would also likely cross-react with the group A streptococcal protein.

#### Example 5

25

**Alignment of Predicted Amino Acid Sequences of the Sp36 Homologs from group A and B Streptococci With Pneumococcal Sp36**

The predicted amino acid sequences from the Sp36 genes from group A and group B streptococci and *S. pneumoniae* were aligned using the Clustal algorithm in a DNASTar Computer package (DNASTar, Inc., Madison, WI). Amino acids that match those encoded by the pneumococcal gene are boxed in Figure 2 (showing the results of the alignment). Gaps introduced in the sequence by the alignment process are indicated by dashed lines.

### Example 6

#### Percentage Sequence Identity Between Homologs of Sp36

The Sp36 amino acid sequence from pneumococci is 25.6% identical to the predicted amino acid sequence of the homologous gene of group B streptococci and 25.1% and 12.6% identical to the deduced sequences of the two genes from group A streptococci. Furthermore, the group B homolog is 97.7% and 11.6% identical to the first (GAS36) and second (GAS36(2)) homologs from group A streptococci, respectively. These experiments indicate that homologous genes to Sp36 from pneumococcus are present in group A and group B streptococci, as well as in *Staphylococcus aureus*. The protein encoded by this gene may therefore perform a similar function in these different organisms. This suggests that a vaccine comprising one or more of these proteins may be broadly protective against these species. These results are summarized in Table 1 which shows the percent identity between the amino acid sequences of Sp36 from pneumococcus strain Norway 4 (serotype 4), group A streptococci Sp36 homolog from an M1 serotype, and group B streptococci Sp36 from strain R268.

Table 1.		Pneumo. Sp36	GAS36	GAS36(2)	GBS36
5	Pneumo. Sp36	100%	25.1%	12.6%	25.6%
	GAS36	---	100%		97.7%
	GAS36(2)	---	---	100%	11.6%
		---	---		100%
	GBS36	---	---		

where     $GAS36 = SEQ\ ID\ NO: 2$   
            $GAS36(2) = SEQ\ ID\ NO: 4$   
            $GBS36 = SEQ\ ID\ NO: 6$